Specification #1: Data Preparation

Version: 0.1 16/SEP/09

Changes

0.1: Initial revision: will define the transformations / processing necessary to take integrated measurements from e.g. fast data processing and prepare them for downstream analysis, including:

- Map calculation and refinement
- Heavy atom substructure calculation

Purpose

This will define the steps which are needed to prepare intensity data for downstream processing in Solve-o-Matic. The definition of the module / interface should allow for the fact that we may be wanting structure factor amplitude estimates or e.g. E values. It will also be necessary to estimate the systematic absences, allow reindexing and assignment of the crystal spacegroup.

Preconditions

The following preconditions are necessary:

- Scaled and merged intensity data are available ideally in MTZ format.
- Correct cell constants are already assigned from data processing.
- If correct spacegroup is known without the correct cell constants, the corresponding reindexing operation should be applied.
- To help with TRUNCATE, provision of the number of residues would be helpful.

Postconditions

The structure factor amplitudes and intensities will be available with the correct spacegroup assigned with the given systematic absences removed and with the predefined setting. N.B. this may not be uniquely correct for downstream processing due to alternative origin definitions.

Error States

To be defined.

Process

FIXME

License

During development this will not be distributed. When it is "finished" the resulting pipeline should be made available to CCP4 using an EDNA framework to provide the interface.